

**Univerzita Karlova v Praze**  
**Přírodovědecká fakulta**

Studijní program: Biologie

Studijní obor: Biologie



**Martina Klincová**

Využití transgenních rostlin při studiu funkce rostlinných genů

Use of transgenic plants in functional analysis of plant genes

Bakalářská práce

Školitel: RNDr. Lukáš Fischer, Ph.D.

Praha, 2021

### **Prehlásenie**

Prehlasujem, že som záverečnú prácu vypracovala samostatne a že som uviedla všetky použité zdroje a literatúru. Táto práca, ani jej podstatná časť, nebola predložená k získaniu iného alebo rovnakého akademického titulu.

V Prahe dňa

.....

Martina Klincová

## **Pod'akovanie**

Týmto by som chcela poďakovať svojmu školiťovi, RNDr. Lukášovi Fischerovi, Ph.D., za trpezlivosť, odborné vedenie, cenné rady a čas, ktorý mi počas vypracovávania práce venoval.

# Abstrakt

## Využitie transgénnych rastlín pri štúdiu funkcie rastlinných génov

Funkčná genomika umožňuje odhaľovať funkcie dosiaľ necharakterizovaných génov. Umožňuje pochopiť nielen funkcie jednotlivých génov, ale aj vzťahy medzi nimi. Získane poznatky je možné využiť aj vo vylepšovaní vlastností rôznych kultúrnych rastlín. Pri znižujúcej sa cene a zvyšujúcej sa účinnosti sekvenovania, sú mnohé rastlinné genómy osekvenované, ale bez priradenej funkcie jednotlivých génov. K charakterizácii týchto génov sa používajú prístupy reverznej genetiky. V prvom rade sú medzi nimi prístupy, medzi ktoré patrí analýza inzerčných mutantov, či použitie RNA interferencie, ktoré inaktivujú alebo znižujú expresiu jednotlivých študovaných génov. Inaktivovať, či cielene upraviť funkciu určitého vybraného génu môžeme aj pomocou miestne špecifických endonukleáz, medzi ktoré patria meganukleázy, nukleázy zinkových prstov, TALEN (transcription activator-like effector nucleases) a CRISPR (clustered regularly interspaced short palindromic repeats). Oproti inaktivačným prístupom sú prístupy ako aktivačná mutagenéza a ektopická expresia, prostredníctvom ktorých naopak expresiu zvýšime. Okrem toho je možné analyzovať expresiu génu a lokalizáciu proteínu pomocou reportérových génov v translačnej a transkripčnej fúzii. V tejto práci zhŕňam metódy reverznej genetiky založené na tvorbe transgénnych rastlín, ale okrajovo spomeniem aj také, ktoré sú založené na iných prístupoch.

**KLÚČOVÉ SLOVÁ:** funkčná genomika, reverzná genetika, identifikácia génov, mutantné rastliny, transgénne rastliny

# **Abstract**

## **Use of transgenic plants in functional analysis of plant genes**

Functional genomics makes it possible to detect the functions of uncharacterized genes. It allows you to understand not only the functions of individual genes, but also the relationships between them. The acquired knowledge can also be used in improving the properties of various cultivated plants. With decreasing cost and increasing sequencing efficiency, many plant genomes are sequenced, but without the associated function of individual genes. Reverse genetics approaches are used to characterize these genes. In the first place, these include approaches that include the analysis of insertional mutants or the use of RNA interference that inactivates or reduces the expression of the individual genes studied. We can also inactivate whether to specifically modify the function of a particular selected gene using site-specific endonucleases, which include meganucleases, zinc finger nucleases, TALEN (transcription activator-like effector nucleases) and CRISPR (clustered regularly interspaced short palindromic repeats). In contrast to inactivation approaches, there are approaches such as activating mutagenesis and ectopic expression, through which we increase the expression. In addition, gene expression and protein localization can be analyzed using reporter genes in translational and transcriptional fusion. In this work I summarize the methods of reverse genetics based on the creation of transgenic plants, but I will also mention those that are based on other approaches.

**KEYWORDS:** functional genomics, reverse genetics, gene identification, mutant plants, transgenic plants

## Obsah

1. Úvod .....	1
2. Reverzná genetika a funkčná genomika .....	2
3. Inaktivácia (loss of function) a cielená editácia .....	4
3.1. Inzerčná mutagenéza .....	4
3.1.1. T-DNA .....	4
3.1.2. TRANSPOZÓNY .....	5
3.2. Miestne špecifické endonukleázy (site-directed nucleases; SDN) .....	5
3.2.1. Meganukleázy .....	6
3.2.2. Nukleázy zinkových prstov (ZINC-FINGER NUCLEASES; ZFN) .....	6
3.2.3. TALEN .....	7
3.2.4. CRISPR/CAS .....	7
3.2.5. Ďalšie využitia miestne špecifických endonukleáz .....	8
3.3. RNA interferencia .....	9
4. GAIN OF FUNCTION .....	12
4.1. Aktivačná inzerčná mutagenéza .....	12
4.2. Ektopická overexpresia .....	13
5. Analýza expresie pomocou reportérových génov .....	15
5.1. Transkripčná fúzia .....	15
5.2. Translačná fúzia .....	15
6. TILLING .....	16
7. Záver .....	18
Zoznam literatúry .....	19

## **Zoznam použitých skratiek**

ABE: adenine base editors

AGO: Argonaut

amiRNA: artificial micro RNA

CAM: crassulacean acid metabolism; metabolizmus kyselín u tučnolistových

CaMV: Cauliflower mosaic virus; vírus mozaiky karfiolu

Cas: CRISPR associated protein

CBE: cytosine base editor

cDNA: complementary DNA

CRISPR: clustered regularly interspaced short palindromic repeats

CRISPRi: CRISPR interference

DHPLC: denaturing high performance liquid chromatography; denaturačná vysokoúčinná kvapalinová chromatografia

DNA: deoxyribonucleic acid; deoxyribonukleová kyselina

DSB: double-strand break; dvojvláknový zlom

EMS: etylmetansulfonát

GFP: green fluorescent protein; zelený fluorescenčný proteín

GUS: glukuronidáza

hpRNA: hairpin RNA

HR: homologická rekombinácia

LUC: luciferáza

miRNA: micro RNA

MMEJ: microhomology-mediated end joining; spájanie koncov sprostredkovaný mikrohomiologiou

NGS: next generation sequencing; sekvenovanie novej generácie

NHEJ: non-homologous end joining; nehomológne spájanie koncov

PAM: protospacer adjacent motif

PCR: polymerase chain reaction; polymerázová reťazová reakcia

pegRNA: prime editing guide RNA

PTGS: post-transcriptional gene silencing; post-transkripčné umlčovanie génov

RdDM: RNA-directed DNA methylation; DNA metylácia riadená RNA

RISC: RNA-induced silencing complex

RNA: ribonucleic acid; ribonukleová kyselina

RNA-Seq: RNA sequencing; sekvenovanie RNA

RNAi: RNA interferencia

RT: reverzná transkriptáza

RVD: variable di-residue

sgRNA: single guide RNA

siRNA: small interfering RNA; malé interferujúce RNA

SMRT: singlemolecule real time

T-DNA: transferová DNA

TALEN: transcription activator-like effector nucleases

TGS: transcriptional gene silencing; transkripčné umlčovanie génov

TILLING – targeting induced local lesions in genomes

VIGS: virus induced gene silencing; vírusom indukované umlčovanie génov

YFP: yellow fluorescent protein; žltý fluorescenčný proteín

ZFN: zinc finger nuclease; nukleáza so zinkovými prstami

ZFP: zinc finger protein; zinkový prst



## 1. Úvod

Transgénnymi rastlinami označujeme také rastliny, ktorých DNA je geneticky modifikovaná, čiže je do nej vložená cudzia, exogénna DNA. Táto exogénna sekvencia je do rastliny vložená transformáciou, najbežnejšie pomocou baktérie *Agrobacterium tumefaciens* alebo biolistickou metódou nastrelení. Biolistická metóda „gene gun“ používa mikroskopické zlaté častice. Množstvo takýchto častíc je obalených DNA s génom, ktorý chceme vložiť do rastliny, a následne nastrelených tlakom do rastlinných buniek. Keď sa guľička dostane do jadra, prípadne plastidu, dôjde k inkorporácii génov do genómu rastliny (Joshi et al., 2011). Druhou metódou je transformácia pomocou agrobaktérie. Je to prirodzená cesta transgenózy. Ide o vloženie úseku DNA do poranenej rastlinnej bunky prostredníctvom aparátu agrobaktérie, ktorá má k tomuto prenosu vyvinutý tzv. Ti plazmid. Gény pre určité enzýmy nachádzajúce sa na tzv. T-DNA Ti plazmidu, môžeme nahradiť vlastnými, a teda preniesť agrobaktériou do rastlinného genómu (Gelvin, 2003).

Význam transgénnych rastlín vo výskume je pri štúdiu funkcie jednotlivých rastlinných génov. Po osekvenovaní prvých genómov modelových rastlín *Arabidopsis thaliana* (Arabidopsis Genome Initiative, 2000) a *Oryza sativa* (International Rice Genome Sequencing Project, 2005) došlo k rozmachu sekvenovania pomocou metód sekvenovania druhej a tretej generácie. Momentálne je osekvenované nesmierne množstvo genómov rôznych organizmov, vrátane človeka (International Human Genome Sequencing Consortium, 2004). Vedci získali množstvo génov, ktorých úlohy sú stále neznáme. Funkcie a prejavy študovaných génov sledujeme pomocou prístupov reverznej genetiky.

## 2. Reverzná genetika a funkčná genomika

Klasická, priama genetika (angl. „forward genetics“) spočíva v identifikácii génov, ktoré sú zodpovedné za určité vlastnosti alebo znaky organizmu. To znamená, že známy je fenotyp a hľadá sa zodpovedajúca sekvencia DNA, čiže postupujeme od fenotypu ku genotypu. Reverzná genetika (angl. „reverse genetics“) však funguje opačným smerom. Pri reverznej genetike postupujeme od genotypu k fenotypu. Ide o rôzne modifikácie určitého študovaného génu, ktorých následky sa prejavujú vo fenotype, a tým umožnia objasniť úlohu tohto génu. Hľadajú sa funkcie už poznaných génových sekvencií. Prístupmi reverznej genetiky zisťujeme, aké sú funkcie a prejavy týchto génov. Pomocou mutagenézy sa pozmení ich sekvencia alebo expresia a sleduje sa, ako sa tieto mutácie prejavujú. Existuje mnoho metód reverznej genetiky. Gén môžeme napríklad inaktivovať, aktivovať na inom mieste, prípadne pozmeniť jeho funkciu.

Na rozdiel od genetiky, ktorá skúma funkciu a zloženie jednotlivých génov, sa genomika venuje všetkým génom v organizme a ich vzťahom, za účelom identifikácie ich vplyvu na organizmus. Má niekoľko odvetví. Jedným z nich je funkčná genomika. V rastlinnom výskume začala byť využiteľná po prvých osekvenovaných rastlinných genómoch (*Arabidopsis* Genome Initiative, 2000; International Rice Genome Sequencing Project, 2005). Cieľom funkčnej genomiky je určiť, ako jednotlivé zložky genómu organizmu spolupracujú na vytvorení konkrétneho fenotypu. Typickými úrovňami štúdia funkčnej genomiky je transkriptomika a proteomika. Transkriptomika je vedecká disciplína, ktorá študuje transkriptóm, čo označuje súbor všetkých RNA nachádzajúcich sa v bunke. Proteomika je oblasť skúmajúca proteíny a vzťahy medzi nimi. Poznatky získané metódami funkčnej genomiky majú dôležitý význam nie len vo výskume rastlín, ale aj v poľnohospodárstve a medicíne.

Poradie jednotlivých nukleotidov v sekvencii DNA určujeme procesom nazývaným sekvenovanie. Ak chceme študovať funkcie jednotlivých génov, musíme v prvom rade poznať primárnu štruktúru nukleovej kyseliny. Je mnoho spôsobov, ktorými tieto gény identifikujeme. Sangerova metóda patrí k najstarším (Sanger & Coulson, 1975; Sanger et al., 1977). Hoci je táto metóda používaná dodnes, postupne je nahradzovaná novšími a efektívnejšími metódami sekvenovania novej generácie (next generation sequencing; NGS). Pri prístupoch druhej generácie paralelne sekvenujeme veľké množstvo malých fragmentov genómu, ktoré sa potom „poskladajú“ a vytvoria celogenómovú sekvenciu DNA. Patrí sem napríklad sekvenovanie na základe ligácie oligonukleotidov (SOLiD) a sekvenovanie pomocou syntézy (Illumina, 454

pyrosekvenovanie). Tieto metódy umožňovali rýchlejšie a lacnejšie získavanie sekvencií a preto sa ľahko dostali do popredia. Pri metódach tretej generácie sa genómová DNA nemusí amplifikovať a nepoužíva sa ani fragmentácia na menšie úseky. Sem radíme SMRT (singlemolecule real time) a nanopor (Bleidorn, 2015).

Metódy NGS sa používajú aj pri analýze transkriptomov, tzv. RNA-Seq (RNA sequencing). Používajú sa rôzne metódy NGS na detekciu celkového stavu transkriptov v bunke. Napriek tomu, že ide o analýzu RNA, sekvenuje sa cDNA. Sú však prípady, keď sa sekvenuje priamo RNA (Cloonan et al., 2008). Ďalšie využitie NGS je pri výskume methylomu, čiže detekciu methylácie DNA (Lister et al., 2008). Na analýzu špecifických mutácií sa používa cielené sekvenovanie, pri ktorom skúmame len určitú vybranú časť genómu (Boutigny et al., 2019).

### 3. Inaktivácia (LOSS OF FUNCTION)

#### 3.1. Inzerčná mutagenéza

Jednou z ciest, ako vyradiť funkciu génov je použitie inzerčnej mutagenézy. Inzerčná mutagenéza je založená na vložení transferovej DNA (T-DNA) alebo transponovateľných elementov do genómu rastliny. Keď sa táto cudzia sekvencia vloží do génu, vyradí, prípadne pozmení funkciu daného génu. Integrácia T-DNA a transpozónov tvorí mutácie v genóme náhodne. Pretože je vkladaná sekvencia známa, inzert poslúži ako marker, podľa ktorého identifikujeme miesto vloženia. Presné umiestnenie mutácie je určené sekvenovaním DNA (Krysan et al., 1996). Vytvorením populácií mutantov bolo možné identifikovať mutácie vo väčšine génov modelových druhov *Arabidopsis thaliana* (Galbiati et al., 2000) a *Oryza sativa* (Jeon et al., 2000). Nevýhodou tohto prístupu je, že inzerty v požadovanom gène napriek tomu nemusia byť v kolekciách dostupné (Krysan et al., 1996).

##### 3.1.1. T-DNA

T-DNA je časť Ti plazmidu gramnegatívnej baktérie *Agrobacterium tumefaciens*. Gény na T-DNA prirodzene kódujú enzýmy, ktoré syntetizujú fytohormóny a opíny. Tieto gény sú nahradené novými a transformáciou, ktorú zaistí táto baktéria, vložené do rastliny (Gelvin, 2003).

Výber miesta integrácie T-DNA v genóme je náhodný (Kim & Gelvin, 2007) vyselektované transformované rastliny majú preferenčne T-DNA integrovanú do miest iniciácie transkripcie alebo terminátora (Sallaud et al., 2004; Li et al., 2006) a taktiež do oblastí, kde je viac AT párov oproti GC párom (Brunaud et al., 2002), teda v miestach, kde zostáva s väčšou pravdepodobnosťou aktívny vnášaný selekčný gén (Kim & Gelvin, 2007).

Pre použitie v reverznej genetike je nutné vytváranie veľkých charakterizovaných mutantných kolekcií a ich analýza. Existuje množstvo mutantných línií rôznych rastlinných druhov. Mutantní linie z týchto zbierok je možné vyhľadať online a sú cenovo dostupné. V *Arabidopsis thaliana* patrí k najznámejším GABI-Kat (Kleinboelting et al., 2017), SALK, SAIL, prípadne WISC (Sessions et al., 2002; Woody et al., 2007; O'Malley et al., 2015). V ryži je to napríklad TRIM (Hsing et al., 2007). U rastliny *Brachypodium distachyon* zas kolekcia od WRRC (Bragg et al., 2012).

### 3.1.2. Transpozóny

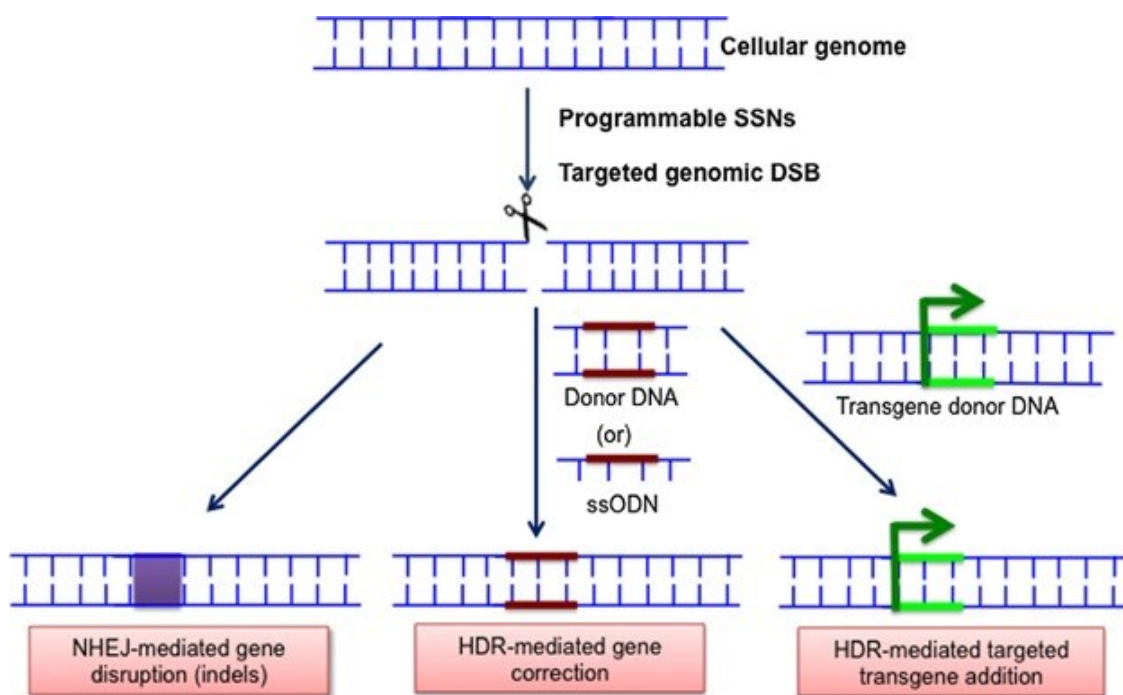
Klasické transpozóny sú úseky DNA, ktoré sú schopné pohybovať sa v genóme, teda preskakovať z miesta na miesto. Prirodzene sú DNA transpozóny umlčané a neaktívne (Muñoz-López & García-Pérez, 2010). Endogénne transpozóny sú prirodzene sa vyskytujúce mutagény u všetkých rastlinných druhov. Najznámejšie sú u kukurice: *Ac/Ds* (McClintock, 1947) a *Spm/En* (Peterson, 1970). Tieto endogénne transpozóny sa využívajú v heterologických systémoch ako mutagény. Pre mutagenézu sa vždy používajú transpozóny z iného druhu, aby sa dala kontrolovať ich aktivita a aby sa nové inzercie dali jednoducho identifikovať (Parinov et al., 1999; Speulman et al., 2000). Schopnosť niektorých transpozónov je preferenčne sa inzerovať do tandemov alebo génových rodín, čím môžu vytvárať inzerčné mutácie všetkých ich členov (Krysan et al., 1999). Transponovateľné elementy sa preferenčne vkladajú do kódujúcich oblastí genómu (Miyao et al., 2003; Kolesnik et al., 2004). Transpozóny, na rozdiel od T-DNA, je možné z génu spätne vystrihnúť, a tak overiť, či pozorované zmeny fenotypu boli skutočne spôsobené inzerciou transpozónu do určitého génu.

### 3.2. Miestne špecifické endonukleázy (site-directed nucleases; SDN)

Mutácie pomocou T-DNA a transpozónov sú v genóme rozmiestnené náhodne a chýba im špecifita, preto sa využívajú spomínané charakterizované kolekcie mutantov. Pomocou miestne špecifických nukleáz môžeme inaktivovať predom zvolený určitý gén. Patria medzi nich nukleázy zinkových prstov (ZFN), TALEN (transcription activator-like effector nucleases), meganukleázy a CRISPR/Cas (clustered regularly interspread short palindromic repeats/CRISPR-associated protein) (Khandagale & Nadaf, 2016). Tieto nukleázy môžeme navrhnúť alebo upraviť pre vybrané miesto štiepenia a následne, vytvorenou nukleázou, v špecifickom lokuse spraviť dvojvláknový zlom (double-strand break; DSB). Vzniknuté dvojvláknové zlomy sú následne opravené bunkovými mechanizmami (Obr.1). Mechanizmy opravujúce dvojvláknové zlomy sú buď homologická rekombinácia (HR), nehomologické spájanie voľných koncov DNA (nonhomologous end joining; NHEJ) (Khanna & Jackson, 2001; Puchta, 2005) alebo mikrohomológiou sprostredkované spájanie voľných koncov DNA (microhomology-mediated end joining; MMEJ) (Lee & Lee, 2007). Každá z nich poskytuje iné možnosti cielennej mutagenézy.

Pri NHEJ či MMEJ sa nepoužíva na opravu homologický templát, preto často vznikajú inzercie alebo delécie (InDel) v mieste zlomu a dochádza tak k vyradeniu funkcie génu (knockout). Oproti tomu, pri oprave pomocou homologickej rekombinácie sa používa homologický úsek DNA, podľa ktorého sa zlom opraví. Podľa charakteru poskytnutej

templátovej DNA je buďto zabezpečená bezchybná oprava, oprava s pripravenou mutáciou, ale môžeme docieľiť aj inzerciu jedného alebo viacerých transgénov do miest dvojvláknového zlomu. Opravy pomocou HR sú menej efektívne, ale prirodzene dochádza k menšej chybovosti (Puchta, 2005). To, či bude DSB opravený pomocou HR alebo NHEJ, závisí aj od časti bunkového cyklu. K NHEJ dochádza častejšie v G2/M fáze, kdežto k HR v S fáze (Bee et al., 2013)



Obr.1: Opravy DSB (prevzaté z (Govindan & Ramalingam, 2016))

### 3.2.1. Meganukleázy

Meganukleázy, často označované ako tzv homing endonukleázy, patria k najstarším sekvenčne špecifickým nástrojom reverznej genetiky. Majú dlhé rozpoznávacie sekvencie. Sú schopné štiepiť DNA sekvencie dlhé až do 40 párov báz. Medzi najznámejšie a najčastejšie používané patria meganukleázy rodiny LAGLIDADG. Prvý krát bola u rastlín použitá nukleáza práve z tejto rodiny, I-SceI, v rastline tabaku *Nicotiana plumbaginifolia*, na vytvorenie dvojvláknového zlomu (Puchta et al., 1993). I-SceI bola použitá aj na DSB a inaktiváciu génu v ryži (Nandy et al., 2015). Ale keďže sú meganukleázy zložité a ich designovanie je časovo náročné, sú často krát nahradzované novšími a efektívnejšími metódami na vytváranie DSB.

### 3.2.2. Nukleázy zinkových prstov (ZINC-FINGER NUCLEASES; ZFN)

Nukleázy zinkových prstov sú umelo vytvorené nukleázy, ktoré fungujú ako diméry. Každý monomér má dve hlavné zložky. Prvá je doména, ktorá sa viaže na DNA a druhá je nukleázová doména restriktnej endonukleázy FokI (Smith et al., 2000). Na DNA sa viažu tri Cys2His2

zinkové prsty (zinc finger protein; ZFP). Jeden zinkový prst rozpoznáva tri nukleotidy. Jeden monomér ZFN je tvorený 3 až 4 ZFP, tým pádom rozpoznáva 9 až 12 nukleotidov na jednom vlákne DNA (Urnov et al., 2010). Medzi monomérmi naviazanými na protiľahlé vlákna DNA, je tzv. spacer sekvencia dlhá zhruba 6 párov báz. Je to miesto, v ktorom pôsobí FokI (Shimizu et al., 2009). FokI je endonukleáza, ktorá je aktívna len vo forme diméru, ktorý sa vytvoril párom ZFN, a vytvorí DSB (Smith et al., 2000).

Lloyd v 2005 presne popísal, ako funguje cielená mutagenéza pomocou ZFN v *Arabidopsis* (Lloyd et al., 2005). Pred ním už tieto nukleázy boli použité v iných organizmoch, po ňom sa začali používať aj v mnohých iných rastlinách.

### **3.2.3. TALEN**

Na konci 2010 sa k meganukleázam a ZFN pridali TALEN, ako nástroje na vytváranie DSB v DNA (Christian et al., 2010; Li et al., 2011). TALEN sú, podobne ako ZFN, umelo vytvorené nukleázy, ktoré fungujú vo forme dimérov. Rovnako je tam FokI restričná endonukleáza. DNA-väzobná doména je však odlišná, tvoria ju TALE proteíny. TALE proteíny sú zložené z 12 – 27 repetitívnych domén o 33 – 35 aminokyselinách. Špecifickosť určujú hlavne aminokyseliny na 12 a 13 pozícii, označované RVD (repeat variable di-residue). Asn-Ile rozpoznáva A, Asn-Gly T, Asn-Asn G a His-Asp rozpoznáva C. Spacerová oblasť je dlhá 15 párov báz (Christian et al., 2010). Dvojvláknové zlomy vytvorené TALENmi boli použité napríklad v ryži (Nishizawa-Yokoi et al., 2016), *Arabidopsis* (Christian et al., 2013), pšenici (Luo et al., 2019) a iných rastlinách.

### **3.2.4. CRISPR/Cas**

CRISPR/Cas funguje ako „imunitný“ systém u baktérií a archeí. Od roku 2013 sa používa na editáciu genómov rastlín, ako je ryža (Tang et al., 2017), sója (Jacobs et al., 2015), *Arabidopsis* (Liu et al., 2015), a mnoho ďalších (Lawrenson et al., 2015). Existuje viac typov CRISPRu. Momentálne najvyužívanejší a najznámejší je typ II, CRISPR/Cas9. Má dve hlavné zložky, čo je sgRNA (single guide RNA) a endonukleáza Cas9. U ZFN a TALEN špecifitu zabezpečovali DNA-väzobné domény proteínov. U CRISPRu ju zabezpečuje v prvom rade sgRNA, konkrétne 20 nukleotidov na 5' konci, v druhom rade väzba Cas9 na trojnukleotidovú PAM sekvenciu (NGG), ktorá sa nachádza v blízkosti cieľového miesta. Na 3' konci sgRNA je 42 nukleotidová vlásenková štruktúra, ktorá viaže endonukleázu Cas9. Hybridizácia sgRNA s cieľovou DNA spôsobí, že endonukleáza Cas9 aktivuje svoje dve nukleázové domény (RuvC a HNH) a produkuje DSB v cieľovej sekvencii. Každá doména Cas9 štiepi jedno vlákno (Jinek et al., 2012, 2013). Na knockout génu v hostiteľskej bunke treba transformovať bunku s génom

pre Cas9 a génom kódujúcim sgRNA, ktorý obsahuje 20 nukleotidov komplementárnych k cieľovému segmentu DNA (Jinek et al., 2013).

Zaujímavý je aj mechanizmus duálnej sgRNA, pri ktorom dve sgRNA naraz „zaútočia“ na jeden gén a tým pádom je knockout spoľahlivejší. Často dochádza k delícii medzi cieľovými miestami jednotlivých sgRNA (Pauwels et al., 2018). Existuje taktiež možnosť zacielenia viacerých génov súčasne, ak sa navrhne viac sgRNA s rôznymi cieľovými sekvenciami (Lowder et al., 2015).

V reverznej genetike bol CRISPR využitý napríklad na charakterizáciu génov spojených s CAM (Crassulacean Acid Metabolism). Knockoutovaním génu pre receptor modrého svetla, fototropín 2, pomocou CRISPRu, sa zistilo, že je dôležitý v CAM dráhe (Liu et al., 2019).

Nedávno sa začal používať aj CRISPR typu V, CRISPR/Cpf1, ako u rastlín, tak aj u iných organizmov. Je špecifickejší ako typ II, jeho riadiaca RNA má s cieľovou DNA komplementárnych 23 nukleotidov a jeho PAM sekvencia je tvorená TTTV, pričom V je adenín, cytozín alebo guanín (Tang et al., 2018).

### **3.2.5. Ďalšie využitia miestne špecifických endonukleáz**

Okrem inaktívácie, môžeme gén pomocou miestne špecifických endonukleáz aj cielene upraviť. Často na základe opravy dvojláknového zlomu pomocou homologickej rekombinácie. Keďže je HR málo efektívna, existujú rôzne spôsoby, ako zvýšiť efektivitu. Jeden prípad je použiť replikony geminiviru, ktoré budú slúžiť ako templát na opravu DSB (Baltes et al., 2014).

Pomocou HR môžeme integrovať do rastlinného genómu aj celý gén, a to na presnom špecifickom mieste. Vnesenie a presná integrácia génu je zabezpečená vložením templátovej DNA, ktorá je ohraničená sekvenciami komplementárnymi so sekvenciami miesta štiepenia v genóme (Svitashev et al., 2015).

Za zmienku stojí CRISPR/dCas9 („dead Cas9“). Ide o upravený neaktívny CRISPR/Cas9, ktorý netvorí DSB a je využívaný v procese tzv. CRISPRi (CRISPR interference). Špecifita regulácie je zabezpečená guide RNA. Naviazaný dCas9 síce DNA neštiepi, ale môže byť prifúzovaný na regulátory transkripcie a tým meniť expresiu génov. Fúzny dCas9 tak môže priamo reprimovať alebo aktivovať transkripciu a taktiež môže modifikovať chromatin. Je taktiež dokázané, že vie zacieliť na viac génov súčasne (Qi et al., 2013).



Pri editácii génov sa využívajú aj editori báz. Sú známe ako editori adenínov (adenine base editors; ABE), tak editori cytozínov (cytosine base editors; CBE). Tieto editori báz (s deaminázovou aktivitou) zabezpečujú substitučné mutácie bez tvorby dvojvláknového zlomu. Funguje to tak, že sa dCas9 sfúzuje s deaminázou a umožnia tak presnú zámenu jednej bázy. Cieľové miesto je rozpoznané pomocou guide RNA, ktorá je v Cas9 a samotnú náhradu zabezpečí deamináza. Výsledkom je potom zámena párov AT za GC a GC za AT, podľa príslušných editorov (Gaudelli et al., 2017; Kang et al., 2018).

K modernejším prístupom patrí tzv. prime editing. Má tri zložky, a to nukleázu štiepiacu DNA podľa guide RNA, reverznú transkriptázu (RT) a pegRNA. Úloha nukleázy je vytvoriť jednovláknový zlom na cieľovom DNA vlákne. Reverzná transkriptáza číta RNA a používa RNA ako templát na vytvorenie komplementárnej DNA. PegRNA slúži ako guide RNA a má 2 časti. Prvá z nich je komplementárna sekvencia viažuca sa na cieľovú DNA a druhou časťou je RNA kódujúca nový edit. PegRNA vytvorí komplex s nukleázou a RT a „prinesie“ ho k cieľovej sekvencii. Nukleáza v tom mieste vytvorí jednovláknový zlom. RT na základe RNA sekvencie z pegRNA dotvorí cieľovú DNA, s tým, že do miesta SSB je vložená nová sekvencia. Needitované vlákno DNA je následne rozštiepené nukleázou a na základe komplementarity dosyntetizované k editovanému vláknu (Anzalone et al., 2019).

### **3.3. RNA interferencia**

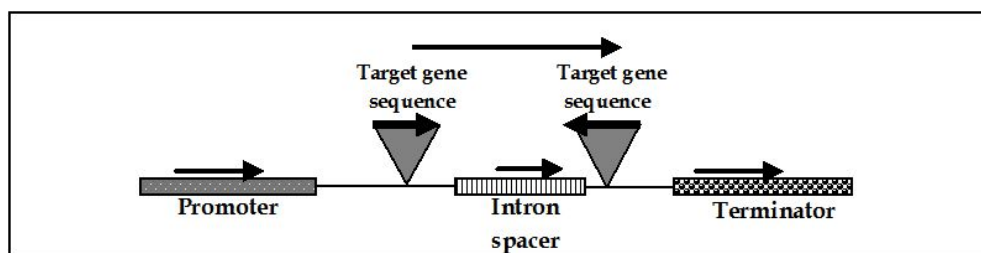
Predchádzajúce metódy vytvárali knockout mutácie, cielene modifikácie sekvencie alebo expresie. Zmeny expresie cieľových génov môžeme dosiahnuť aj RNA interferenciou, vtedy hovoríme o knockdown mutácii, pri ktorej sa gén nevyradí z funkcie úplne, len sa zníži jeho expresia. RNAi môže umlčovať génovú expresiu buď na úrovni transkripcie (transcriptional gene silencing; TGS), alebo na post-transkripčnej úrovni (post-transcriptional gene silencing; PTGS). TGS sa zabezpečuje najmä metyláciou cytozínu v promotore, čo môže vyvolať tvorbu heterochromatínu a modifikácie histónov. V prípade PTGS dochádza najčastejšie k rozštiepeniu mRNA proteínom Argonaut (AGO), prípadne môže dôjsť aj k bloku translácie (Méndez et al., 2015).

RNA interferencia je vyvolaná, keď sa v bunke ocitne dsRNA (double-stranded RNA, dvojvláknová RNA). Táto dsRNA je naštiepaná DICER-like (DCL) enzýmami na 21 – 24 nukleotidov dlhé dvojvláknové úseky, tzv. malé RNA (sRNA). Jedno vlákno tejto malej RNA spolu s proteínom AGO a inými komponentami tvoria RISC (RNA induced silencing complex). Následne sa pomocou RISC, konkrétne vláknu malej RNA, ktorá je jeho súčasťou, na základe

homológie vyberie cieľová molekula RNA, ktorá bude umlčaná (Baulcombe, 2004; Baumberger & Baulcombe, 2005).

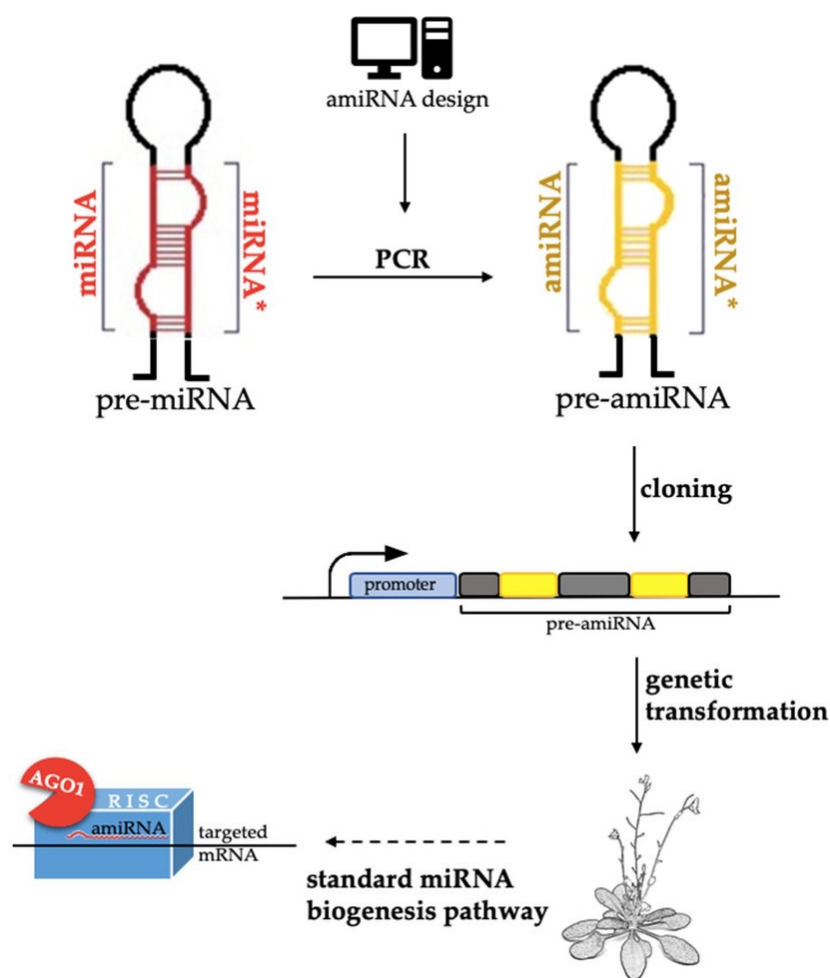
Základom RNAi sú už spomínané malé, nekódujúce RNA. Sú to siRNA (small interfering RNA) a miRNA (microRNA). Si RNA sprostredkovávajú umlčovanie génov aj na úrovni transkripcie, aj na post-transkripčnej úrovni, Kdežto miRNA pôsobia hlavne post-transkripčne (Filipowicz et al., 2005).

Pri štúdiu funkcie génov ich umlčovaním na post-transkripčnej úrovni sa využívajú hlavne hpRNA (hairpinRNA) a amiRNA (artificial microRNA). Konštrukt hpRNA je zložený z dvoch úsekov rovnakých sekvencií cieľového génu v orientácii invertovaných repetícií, medzi ktorými je spacer sekvencia (Obr.2). Po introdukovaní do bunky sa transkripciou takto vytvoreného konštrukt vytvorí hpRNA s dvojvláknovým úsekom tvoreným invertovanými repetíciami. Táto dvojvláknová RNA je v bunke spracovaná na siRNA, ktoré v procese RNA interferencie umlčujú daný gén (Wesley et al., 2001).



Obr.2: Konštrukt hpRNA (prevzaté z Tenea a Burlibasa 2012)

Prekursorami amiRNA sa vytvárajú s prekursorov miRNA. AmiRNA sa navrhuje tak, že sekvencie v miRNA prekuzore sú nahradené sekvenciami komplementárnymi k cieľovým génom (Obr.3). Pomocou enzýmov potrebných v štandardnej dráhe miRNA sa pomocou amiRNA dosiahne umlčanie ktoréhokoľvek génu z genómu organizmu (Warthmann et al., 2008). Môže byť nadesignovaná tak, že umlčí aj viac génov. Je to veľmi presná a efektívna inaktivácia génov (Zhang et al., 2018).



Obr.3.: Dráha a design amiRNA (prevzaté z Wojcik 2020)

Umlčanie génov na úrovni transkripcie sa deje procesom tzv. RNA riadená DNA metylácia (RdDM; RNA directed DNA methylation). Ide o metyláciu cytozínov v promotoroch. Konštrukt sa v tomto prípade navrhuje tak, aby dsRNA bola homologická s oblasťou promotora (Jones et al., 2001). Samotný priebeh RdDM začína štiepením dsRNA na malé RNA. Potom jedno vlákno z malej RNA spolu s proteínom AGO indukuje v promotore de novo metyláciu DNA metyltransferázou (Matzke & Moshier, 2014).

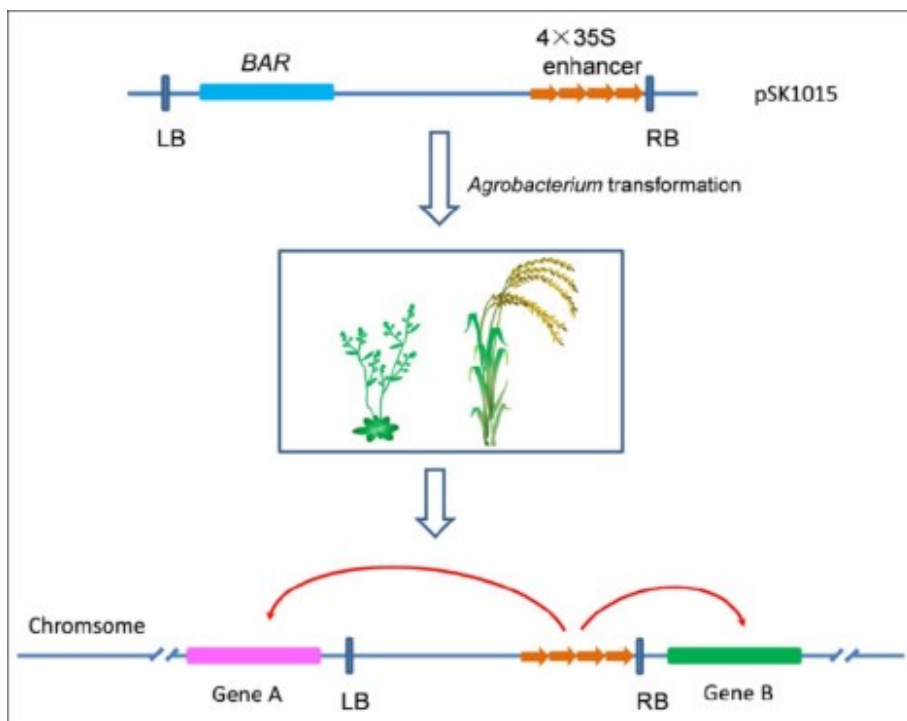
Ako nástroj k štúdiu funkcie rastlinných génov sa taktiež používa umlčanie génu indukované vírusom (virus induced gene silencing; VIGS), nevznikajú pri ňom však transgénne rastliny. Vírusový vektor nesúci gén hostiteľa je vpravený do rastliny pomocou *Agrobacteria*, prípadná je možnosť infekcie po mechanickom poranení listu aj priamo virovým inokulom. RNA interferencia sa spúšťa ako obrana proti vírusovým molekulám, pri čom vzniknú malé RNA aj proti pridanému rastlinnému génu, ktorý je potom post-transkripčne umlčaný, podobne ako pri použití hpRNA (Baulcombe, 1999). VIGS môže byť použitý ako alternatívny prístup u rastlín, ktoré sa nedajú transformovať, napríklad konope (Schachtsiek et al., 2019).

## 4. GAIN OF FUNCTION

Prístupy loss of function, pri ktorých dochádza ku knock-outovaniu génu, majú pri určovaní funkcie génov aj nedostatky. Komplikované je napríklad použitie na funkčne redundantné, či esenciálne gény. A keďže ide o mutácie recesívne, na analýzu potrebujeme homozygotných mutantov (Lodish et al., 2000; Cutler & McCourt, 2005). Ako doplnok a alternatíva k nim sa teda používajú prístupy gain of function, vrátane analýzy aktivačných inzerčných mutantov a ektopickej overexpresie. Aktivačná inzerčná mutagenéza je založená na náhodnej inzercii zosilňovačov transkripcie alebo silných promotorov do genómu rastliny, čím sa zvýši expresia náhodných génov. Tieto nadmerne exprimujúce sa gény sa prejavujú dominantnými fenotypmi (Lodish et al., 2000). Pri ektopickej expresii ide o „zapnutie“ génu tam, kde nie je prirodzene aktívny.

### 4.1. Aktivačná inzerčná mutagenéza

Aktivačná inzerčná mutagenéza je metóda, pri ktorej sa generujú transgénne rastliny pomocou T-DNA vektorov obsahujúcich promotory alebo zosilňovače transkripcie, najčastejšie CaMV (Cauliflower mosaic virus) 35S typu. Tieto inzercie indukujú zvýšenú expresiu susedných génov (Gou & Li, 2011).



Obr.4: Aktivačná inzerčná mutagenéza (prevzaté z Meng et al. 2012)

Príkladom aktivačného konštruktu je vektor pre *Agrobacterium tumefaciens* obsahujúci tetramer CaMV 35S enhancera pri pravej hranici, ktorý sa transformáciou vloží do rastliny

(Obr.4). Takáto T-DNA je náhodne integrovaná do rastlinného genómu a vytvára dominantné mutácie zvýšením expresie génov v jej blízkosti. Vložené zosilňovače aktivujú expresiu génov nachádzajúcich sa v blízkosti inzercie, nezávisle na smere, upstream aj downstream (Kakimoto, 1996).

Dominantné mutanty vzniknuté aktivačnou inzerčnou mutagenézou pomocou T-DNA majú význam nie len v určovaní funkcie génov. Boli napríklad využité aj v priamej genetike k izolovaniu mutantov lepšie odolávajúcich acidite (Bissoli et al., 2020), tolerantných k zvýšenej salinite (Nguyen et al., 2018) a odolných voči nedostatku dôležitých prvkov (Lee et al., 2011).

Sú prípady, keď je použitie T-DNA na aktivačnú inzerciu neefektívne, prípadne rastliny sú netransformovateľné. Vtedy sa používajú na mutagenézu transpozóny (Qu et al., 2008; Davies et al., 2019).

Rovnako ako pri inaktivačných prístupoch, aj pri aktivačných mutantoch máme k dispozícii charakterizované kolekcie (Hsing et al., 2007; Qu et al., 2008).

## **4.2. Ektopická overexpresia**

Do gain of function prístupov reverznej genetiky radíme aj ektopickú expresiu génov. Sú pri nej gény exprimované v častiach, kde prirodzene exprimované nie sú. Môže ísť o gény endogénne (Criqui et al., 2000; Kalyna et al., 2003) aj exogénne, to znamená, že je gén jednej rastliny exprimovaný v inom druhu.

Pri ektopickej expresii sa využívajú hlavne konštitutívne promótori. To sú také, ktoré sú exprimované za všetkých podmienok (Qin et al., 2010). Veľmi často používaný je práve promótor CaMV 35S (Holtorf et al., 1995). Naproti konštitutívnym rozoznávame aj indukibilné, ktoré sa zapínajú iba pri určitých podmienkach či ošetrovaní. Často ide o rôzne stresory, ako je napríklad chlad (Pino et al., 2007).

Heterologická expresia cDNA je prístup, pri ktorom charakterizujeme rastlinné gény v cudzích organizmoch (Cordier et al., 1999). Pomocou rastlinnej cDNA vlozenej do kvasinky sa identifikovali napríklad gény podieľajúce sa na odolnosti voči NaCl (Sentenac et al., 1992).

Častejšie sa pre ektopickú expresiu génov z rôznych rastlín používa modelová rastlina *Arabidopsis thaliana*. Napríklad pri expresii bol-miR390a génov brokolice v *Arabidopsis* sa zistilo, že sú zodpovedné za vývoj, rast a tvorbu laterálnych orgánov (He et al., 2020), TaARF

pšenice za celkový vývoj rastliny (Yao et al., 2009) a domény NB-ARC z viniču za rezistenciu k patogénom (Wen et al., 2015).

## 5. Analýza expresie pomocou reportérových génov

Analyzovať funkcie génov môžeme pomocou reportérových génov, ako sú zelený fluorescenčný proteín (GFP), glukuronidáza (GUS), luciferáza (LUC) (Kavita & Burma, 2008) a žltý fluorescenčný proteín (YFP) (Zhao et al., 2008). Môžeme pomocou nich sledovať expresiu génu, či dokonca lokalizáciu jeho proteínového produktu v rámci bunky. Podľa designu génového konštruktu s reportérovým génom a účelu, rozoznávame transkripčnú fúziu a translačnú fúziu.

### 5.1. Transkripčná fúzia

Pri tzv. transkripčnej fúzii sa analyzuje ako, kde a kedy je gén exprimovaný. Promotor analyzovaného génu sfúzujeme s reportérovým génom a to tak, že reportérový gén bude v downstream smere od promotoru (Obr.6A). Takýto konštrukt je transformáciou vložený do rastliny. Ak je promotor študovaného génu aktívny, tak je syntetizovaný aj daný reportérový gén. Výsledky sledujeme pomocou fluorescenčnej alebo konfokálnej mikroskopie (Lee et al., 2006; Tovar-Aguilar et al., 2019). Z aktivity promotoru sa snažíme pochopiť funkciu génu, čas aktivácie promotoru, určenie pletivovej špecifity (Xiao et al., 2010)



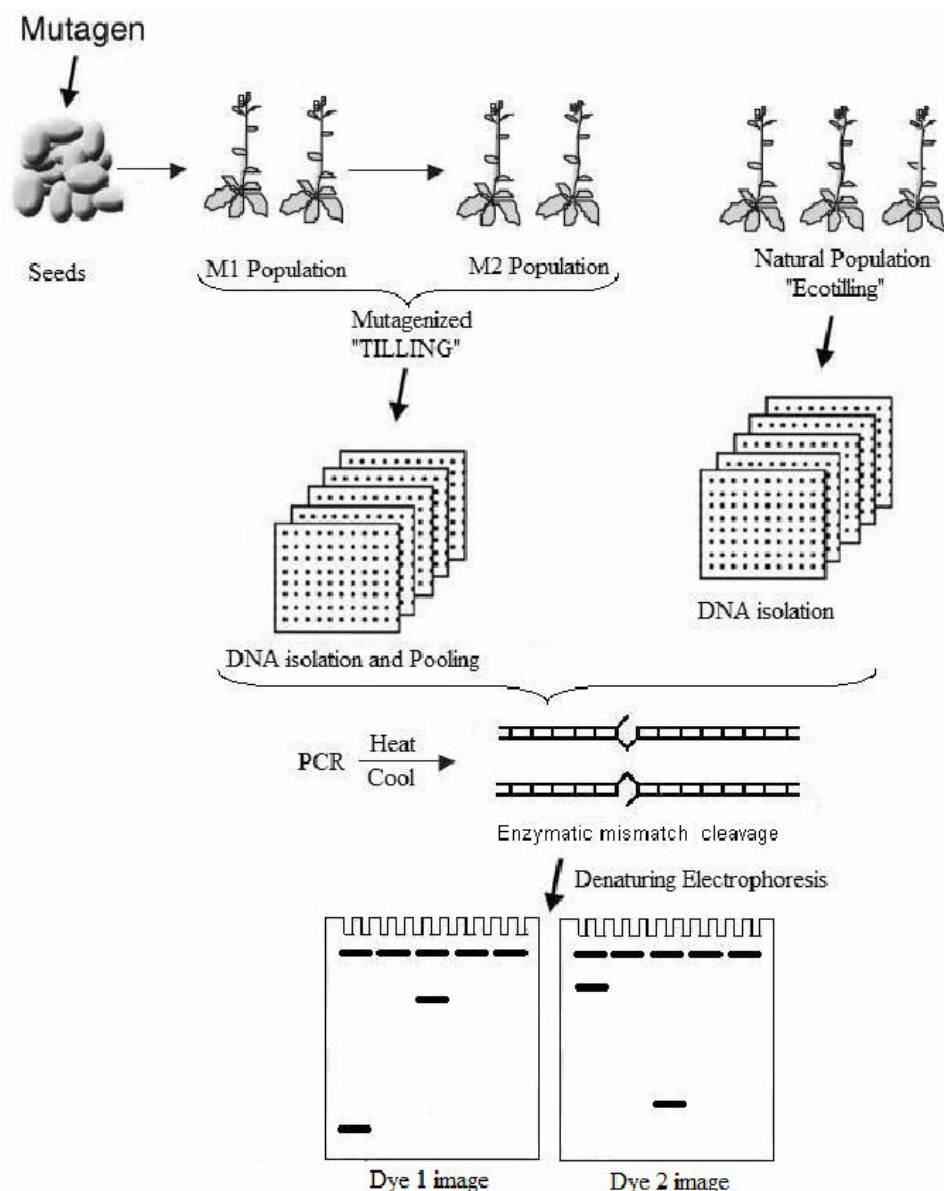
Obr.6: Transkripčná fúzia s GFP (prevzaté z Boulin et al. 2006 - upravené)

### 5.2. Translačná fúzia

Pri translačnej fúzii je reportérová značka (fluorescenčný proteín) fúzovaná so samotným študovaným proteínom (Obr.6B) Primárne je fúzia vytvorená na úrovni DNA a prekladom fúzneho génu vznikne fúzny proteín. Fluorescenčný proteín môže byť prifúzovaný na ktorýkoľvek koniec a aj do stredu proteínu (Deuschle et al., 2005). Ak sa fúzia dobre vytvorí, tak je proteín funkčný a je schopný zastúpiť proteín v mutovaných rastlinách. Môžeme takto zistiť, kde sa študovaný proteín prirodzene vyskytuje a aj sledovať jeho dynamiku, a to pozorujeme konfokálnou alebo fluorescenčnou mikroskopiou (Zhao et al., 2008).

## 6. TILLING

Tilling (Targeted Induced Local Lesions In Genome) je prístup reverznej genetiky, pri ktorom nevznikajú transgénne rastliny. Tento prístup má široké spektrum využitia (Till et al., 2004), ale s rozvojom editácií genómu stráca na význame. Ide o cieľnú identifikáciu bodových mutácií v populácii mutantov. Na indukciu mutácií sa používa klasická necieľená chemická mutagenéza, najčastejšie pomocou EMS (etylmetánsulfonát).



Obr.5: Tilling postup (prevzaté z Şimşek a Aka Kacar 2011)

Samotný proces (Obr.5) začína ošetrovaním semienok M1 generácie mutagénom. Mutácie detekujeme až u M2 rastlín (druhá generácia mutantov). Pri prvom úspešnom použití metódy TILLING v rastlinách (McCallum et al., 2000) sa k detekcii mutácii používala DHPLC (denaturačná vysokoúčinná kvapalinová chromatografia). Tá bola ale nahradená novšími



metódami, najčastejšie sa na detekciu používa špecifický enzým CEL I. CEL I je endonukleáza pochádzajúca zo zeleru, ktorá rozpoznáva chybné párovanie báz, a štiepi teda sekvencie DNA v mieste heteroduplexov. Samotná detekcia mutácie spočíva v tom, že po amplifikácii vybraného génu pomocou PCR sa produkty reakcie denaturujú a spätne renaturujú. Ak je prítomná DNA s mutáciou v danom géne, v mieste mutácie sa vytvorí spomínaný heteroduplex, ktorý je štiepený pomocou CEL I. Mutácie sa potom identifikujú elektroforézou na LI-COR géle, ktorý určí polohu mutácie (McCallum et al., 2000; Till, 2004).

EcoTILLING je prístup odvodený z TILLINGu. Využíva sa na výskum prirodzených jednonukleotidových polymorfizmov. Postup je, až na chemickú mutagenézu, rovnaký. Detekujeme pri ňom prirodzené mutácie. Bol vyvinutý u rastlín, konkrétne *Arabidopsisu*, ale vzhľadom k masívnemu celogenómovému sekvenovaniu rady ekotypov už nie je u tejto rastliny využívaný (Comai et al., 2004).

## 7. Záver

Existuje veľa prístupov na identifikáciu génov pomocou transgénnych rastlín. Každý je osobitý a používa sa v určitých prípadoch. RNA interferencia pomocou dvojvláknových RNA nevyradí gén z funkcie úplne, ale iba stlmí expresiu, hovoríme teda, že ide o knockdown. Ku génovému knock outu dochádza najčastejšie pomocou inzerčnej mutagenézy, teda vloženiu transferovej DNA alebo transponovateľného elementu do génu. Ako doplnok používame prístupy gain of function, ako je ektopická expresia, či analýza aktivačných inzerčných mutantov. Veľké využitie majú aj miestne špecifické endonukleázy, medzi ktoré patria meganukleázy, nukleázy so zinkovými prstami, TALE nukleázy, či CRISPR/Cas, pomocou ktorých vieme zabezpečiť génový knock out, ale niekedy vieme aj špecificky editovať vybraný gén. Nápomocná môže byť aj analýza génovej expresie pomocou reportérových génov, pomocou ktorej sa určuje miesto expresie génu, ako aj lokalizácia proteínov a ich interakcie.

Na analýzu funkcie génov nám nestačí poznať funkcie génov, ale aj vzťahy medzi nimi. Identifikácia génov má obrovské využitie, a to nie len v základnom výskume. Znalosť funkcií génov nám umožňuje zabezpečiť požadované vlastnosti jednotlivých rastlín. Môžeme napríklad zabezpečiť rezistenciu voči patogénom a rôznym abiotickým stresorom, zlepšiť úrodu kultúrnych rastlín, prípadne vytvoriť rastliny s lepšími estetickými vlastnosťami.

## Zoznam literatúry

- Anzalone A.V., Randolph P.B., Davis J.R., Sousa A.A., Koblan L.W., Levy J.M., Chen P.J., Wilson C., Newby G.A., Raguram A. & Liu D.R. 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. **576**: 149–157.
- Arabidopsis Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. **408**: 796–815.
- Baltes N.J., Gil-Humanes J., Cermak T., Atkins P.A. & Voytas D.F. 2014. DNA replicons for plant genome engineering. *Plant Cell*. **26**: 151–163.
- Baulcombe D. 2004. RNA silencing in plants. *Nature*. **431**: 356–363.
- Baulcombe D.C. 1999. Fast forward genetics based on virus-induced gene silencing. *Current Opinion in Plant Biology*. **2**: 109–113.
- Baumberger N. & Baulcombe D.C. 2005. *Arabidopsis* ARGONAUTE1 is an RNA slicer that selectively recruits microRNAs and short interfering RNAs. *Proceedings of the National Academy of Sciences of the United States of America*. **102**: 11928–33.
- Bee L., Fabris S., Cherubini R., Mognato M. & Celotti L. 2013. The efficiency of homologous recombination and non-homologous end joining systems in repairing double-strand breaks during cell cycle progression. *PLoS One*. **8**: e69061.
- Bissoli G., Muñoz-Bertomeu J., Bueso E., Sayas E., Vilcara E.A., Felipe A., Niñoles R., Rubio L., Fernández J.A. & Serrano R. 2020. An *Arabidopsis* Mutant Over-Expressing Subtilase SBT4.13 Uncovers the Role of Oxidative Stress in the Inhibition of Growth by Intracellular Acidification. *Int J Mol Sci*. **21**: .
- Bleidorn C. 2015. Third generation sequencing: Technology and its potential impact on evolutionary biodiversity research. *Systematics and Biodiversity*. **14**: 1–8.
- Boulin T., Etchberger J.F. & Hobert O. 2006. Reporter gene fusions. *WormBook*,
- Boutigny A.-L., Barranger A., De Boissésou C., Blanchard Y. & Rolland M. 2019. Targeted Next Generation Sequencing to study insert stability in genetically modified plants. *Scientific Reports*. **9**: 2308.
- Bragg J.N., Wu J., Gordon S.P., Guttman M.E., Thilmony R., Lazo G.R., Gu Y.Q. & Vogel J.P. 2012. Generation and Characterization of the Western Regional Research Center *Brachypodium* T-DNA Insertional Mutant Collection. *PLOS ONE*. **7**: e41916.
- Brunaud V., Balzergue S., Dubreucq B., Aubourg S., Samson F., Chauvin S., Bechtold N., Cruaud C., DeRose R., Pelletier G., Lepiniec L., Caboche M. & Lecharny A. 2002. T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. *EMBO Rep*. **3**: 1152–1157.
- Christian M., Cermak T., Doyle E.L., Schmidt C., Zhang F., Hummel A., Bogdanove A.J. & Voytas D.F. 2010. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics*. **186**: 757–761.
- Christian M., Qi Y., Zhang Y. & Voytas D.F. 2013. Targeted Mutagenesis of *Arabidopsis thaliana* Using Engineered TAL Effector Nucleases. *G3 Genes|Genomes|Genetics*. **3**: 1697–1705.

- Cloonan N., Forrest A.R.R., Kolle G., Gardiner B.B.A., Faulkner G.J., Brown M.K., Taylor D.F., Steptoe A.L., Wani S., Bethel G., Robertson A.J., Perkins A.C., Bruce S.J., Lee C.C., Ranade S.S., Peckham H.E., Manning J.M., McKernan K.J. & Grimmond S.M. 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods*. **5**: 613–619.
- Comai L., Young K., Till B.J., Reynolds S.H., Greene E.A., Codomo C.A., Enns L.C., Johnson J.E., Burtner C., Odden A.R. & Henikoff S. 2004. Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J*. **37**: 778–786.
- Cordier H., Karst F. & Bergès T. 1999. Heterologous expression in *Saccharomyces cerevisiae* of an *Arabidopsis thaliana* cDNA encoding mevalonate diphosphate decarboxylase. *Plant Mol Biol*. **39**: 953–967.
- Criqui M.C., Parmentier Y., Derevier A., Shen W.-H., Dong A. & Genschik P. 2000. Cell cycle-dependent proteolysis and ectopic overexpression of cyclin B1 in tobacco BY2 cells. *The Plant Journal*. **24**: 763–773.
- Cutler S. & McCourt P. 2005. Dude, Where's My Phenotype? Dealing with Redundancy in Signaling Networks. *Plant Physiol*. **138**: 558–559.
- Davies J.P., Reddy V.S., Liu X.L., Reddy A.S., Ainley W.M., Folkerts O., Marri P., Jiang K. & Wagner D.R. 2019. Development of an activation tagging system for maize. *Plant Direct*. **3**: e00118.
- Deuschle K., Okumoto S., Fehr M., Looger L.L., Kozhukh L. & Frommer W.B. 2005. Construction and optimization of a family of genetically encoded metabolite sensors by semirational protein engineering. *Protein Sci*. **14**: 2304–2314.
- Filipowicz W., Jaskiewicz L., Kolb F.A. & Pillai R.S. 2005. Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr Opin Struct Biol*. **15**: 331–341.
- Galbiati M., Moreno M.A., Nadzan G., Zourelidou M. & Dellaporta S.L. 2000. Large-scale T-DNA mutagenesis in *Arabidopsis* for functional genomic analysis. *Funct Integr Genomics*. **1**: 25–34.
- Gaudelli N.M., Komor A.C., Rees H.A., Packer M.S., Badran A.H., Bryson D.I. & Liu D.R. 2017. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. **551**: 464–471.
- Gelvin S.B. 2003. *Agrobacterium*-Mediated Plant Transformation: the Biology behind the “Gene-Jockeying” Tool. *Microbiol Mol Biol Rev*. **67**: 16–37.
- Gou X. & Li J. 2011. Activation Tagging. pp. 117–133. In: Wang Z.-Y. & Yang Z. (eds), *Plant Signalling Networks*, Humana Press, Totowa, NJ.
- Govindan G. & Ramalingam S. 2016. Programmable Site-Specific Nucleases for Targeted Genome Engineering in Higher Eukaryotes. *Journal of Cellular Physiology*. **231**: 2380–2392.
- He L., Wang Y., Jia Y., Yang Y., Han X., Yuan J., Li L., Li P., Chen C., Song W., Liu M., Li H. & Wang C. 2020. Ectopic overexpression of *bol-miR390a* from broccoli (*B. oleracea* L var. *italica*) increases lateral branches in *Arabidopsis*. *Plant Growth Regul*. **92**: 547–558.
- Holtorf S., Apel K. & Bohlmann H. 1995. Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*. *Plant Mol Biol*. **29**: 637–646.

- Hsing Y.-I., Chern C.-G., Fan M.-J., Lu P.-C., Chen K.-T., Lo S.-F., Sun P.-K., Ho S.-L., Lee K.-W., Wang Y.-C., Huang W.-L., Ko S.-S., Chen S., Chen J.-L., Chung C.-I., Lin Y.-C., Hour A.-L., Wang Y.-W., Chang Y.-C., Tsai M.-W., Lin Y.-S., Chen Y.-C., Yen H.-M., Li C.-P., Wey C.-K., Tseng C.-S., Lai M.-H., Huang S.-C., Chen L.-J. & Yu S.-M. 2007. A rice gene activation/knockout mutant resource for high throughput functional genomics. *Plant Mol Biol.* **63**: 351–364.
- International Human Genome Sequencing Consortium 2004. Finishing the euchromatic sequence of the human genome. *Nature.* **431**: 931–945.
- International Rice Genome Sequencing Project 2005. The map-based sequence of the rice genome. *Nature.* **436**: 793–800.
- Jacobs T., LaFayette P., Schmitz R. & Parrott W. 2015. Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnology.* **15**: .
- Jeon J.S., Lee S., Jung K.H., Jun S.H., Jeong D.H., Lee J., Kim C., Jang S., Yang K., Nam J., An K., Han M.J., Sung R.J., Choi H.S., Yu J.H., Choi J.H., Cho S.Y., Cha S.S., Kim S.I. & An G. 2000. T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* **22**: 561–570.
- Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A. & Charpentier E. 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science.* **337**: 816–821.
- Jinek M., East A., Cheng A., Lin S., Ma E. & Doudna J. 2013. RNA-programmed genome editing in human cells. *eLife.* **2**: e00471.
- Jones L., Ratcliff F. & Baulcombe D.F. 2001. RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. *Curr. Biol.* **11**: 747–757.
- Joshi M., Mishra A. & Jha B. 2011. Efficient genetic transformation of *Jatropha curcas* L. by microprojectile bombardment using embryo axes. *Industrial Crops and Products.* **33**: 67–77.
- Kakimoto T. 1996. CKI1, a Histidine Kinase Homolog Implicated in Cytokinin Signal Transduction. *Science.* **274**: 982–985.
- Kalyna M., Lopato S. & Barta A. 2003. Ectopic Expression of atRSZ33 Reveals Its Function in Splicing and Causes Pleiotropic Changes in Development. *Mol Biol Cell.* **14**: 3565–3577.
- Kang B.-C., Yun J.-Y., Kim S.-T., Shin Y., Ryu J., Choi M., Woo J.W. & Kim J.-S. 2018. Precision genome engineering through adenine base editing in plants. *Nature Plants.* **4**: 427–431.
- Kavita P. & Burma P.K. 2008. A comparative analysis of green fluorescent protein and beta-glucuronidase protein-encoding genes as a reporter system for studying the temporal expression profiles of promoters. *J Biosci.* **33**: 337–343.
- Khandagale K. & Nadaf A. 2016. Genome editing for targeted improvement of plants. *Plant Biotechnol Rep.* **10**: 327–343.
- Khanna K. & Jackson S. 2001. Khanna KK & Jackson SP. DNA double-strand breaks: Signaling, repair and the cancer connection. *Nature Genet.* **27**: 247–254. *Nature genetics.* **27**: 247–54.

- Kim S.-I. & Gelvin S.B. 2007. Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *The Plant Journal*. **51**: 779–791.
- Kleinboelting N., Huet G. & Weisshaar B. 2017. Enhancing the GABI-Kat *Arabidopsis thaliana* T-DNA Insertion Mutant Database by Incorporating Araport11 Annotation. *Plant Cell Physiol.* **58**: .
- Kolesnik T., Szeverenyi I., Bachmann D., Kumar C.S., Jiang S., Ramamoorthy R., Cai M., Ma Z.G., Sundaresan V. & Ramachandran S. 2004. Establishing an efficient Ac/Ds tagging system in rice: large-scale analysis of Ds flanking sequences. *The Plant Journal*. **37**: 301–314.
- Krysan P.J., Young J.C. & Sussman M.R. 1999. T-DNA as an insertional mutagen in *Arabidopsis*. *Plant Cell*. **11**: 2283–2290.
- Krysan P.J., Young J.C., Tax F. & Sussman M.R. 1996. Identification of transferred DNA insertions within *Arabidopsis* genes involved in signal transduction and ion transport. *PNAS*. **93**: 8145–8150.
- Lawrenson T., Shorinola O., Stacey N., Li C., Ostergaard L., Patron N., Uauy C. & Harwood W. 2015. Induction of targeted, heritable mutations in barley and *Brassica oleracea* using RNA-guided Cas9 nuclease. *Genome Biology*. **16**: .
- Lee J.-Y., Colinas J., Wang J.Y., Mace D., Ohler U. & Benfey P.N. 2006. Transcriptional and posttranscriptional regulation of transcription factor expression in *Arabidopsis* roots. *Proc Natl Acad Sci U S A*. **103**: 6055–6060.
- Lee K. & Lee S.E. 2007. *Saccharomyces cerevisiae* Sae2- and Tel1-Dependent Single-Strand DNA Formation at DNA Break Promotes Microhomology-Mediated End Joining. *Genetics*. **176**: 2003–2014.
- Lee S., Persson D.P., Hansen T.H., Husted S., Schjoerring J.K., Kim Y.-S., Jeon U.S., Kim Y.-K., Kakei Y., Masuda H., Nishizawa N.K. & An G. 2011. Bio-available zinc in rice seeds is increased by activation tagging of nicotianamine synthase. *Plant Biotechnology Journal*. **9**: 865–873.
- Li T., Huang S., Jiang W.Z., Wright D., Spalding M.H., Weeks D.P. & Yang B. 2011. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Research*. **39**: 359–372.
- Li Y., Rosso M.G., Ülker B. & Weisshaar B. 2006. Analysis of T-DNA insertion site distribution patterns in *Arabidopsis thaliana* reveals special features of genes without insertions. *Genomics*. **87**: 645–652.
- Lister R., O'Malley R.C., Tonti-Filippini J., Gregory B.D., Berry C.C., Millar A.H. & Ecker J.R. 2008. Highly Integrated Single-Base Resolution Maps of the Epigenome in *Arabidopsis*. *Cell*. **133**: 523–536.
- Liu D., Chen M., Mendoza B., Cheng H., Hu R., Li L., Trinh C.T., Tuskan G.A. & Yang X. 2019. CRISPR/Cas9-mediated targeted mutagenesis for functional genomics research of crassulacean acid metabolism plants. *Journal of Experimental Botany*. **70**: 6621–6629.
- Liu W., Zhu X., Lei M., Xia Q., Botella J., Zhu J. & Mao Y. 2015. A detailed procedure for CRISPR/Cas9-mediated gene editing in *Arabidopsis thaliana*. *Science Bulletin*. **60**: .
- Lloyd A., Plaisier C.L., Carroll D. & Drews G.N. 2005. Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 2232–2237.

- Lodish H., Berk A., Zipursky S.L., Matsudaira P., Baltimore D. & Darnell J. 2000. Mutations: Types and Causes. *Molecular Cell Biology*. 4th edition. .
- Lowder L.G., Zhang D., Baltes N.J., Paul J.W., Tang X., Zheng X., Voytas D.F., Hsieh T.-F., Zhang Y. & Qi Y. 2015. A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiology*. **169**: 971–985.
- Luo M., Li H., Chakraborty S., Morbitzer R., Rinaldo A., Upadhyaya N., Bhatt D., Louis S., Richardson T., Lahaye T. & Ayliffe M. 2019. Efficient TALEN-mediated gene editing in wheat. *Plant Biotechnology Journal*. **17**: 2026–2028.
- Matzke M.A. & Mosher R.A. 2014. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics*. **15**: 394–408.
- McCallum C.M., Comai L., Greene E.A. & Henikoff S. 2000. Targeted screening for induced mutations. *Nat Biotechnol*. **18**: 455–457.
- McClintock B. 1947. Mutable loci in maize. *Mutable loci in maize*. .
- Méndez C., Ahlenstiel C.L. & Kelleher A.D. 2015. Post-transcriptional gene silencing, transcriptional gene silencing and human immunodeficiency virus. *World J Virol*. **4**: 219–244.
- Meng Y., Hongyu L., Zhao T., Zhang C., Lin C. & Liu B. 2012. Transgenic Plants as Gene-Discovery Tools.
- Miyao A., Tanaka K., Murata K., Sawaki H., Takeda S., Abe K., Shinozuka Y., Onosato K. & Hirochika H. 2003. Target Site Specificity of the Tos17 Retrotransposon Shows a Preference for Insertion within Genes and against Insertion in Retrotransposon-Rich Regions of the Genome. *The Plant Cell*. **15**: 1771–1780.
- Muñoz-López M. & García-Pérez J.L. 2010. DNA Transposons: Nature and Applications in Genomics. *Curr Genomics*. **11**: 115–128.
- Nandy S., Zhao S., Pathak B.P., Manoharan M. & Srivastava V. 2015. Gene stacking in plant cell using recombinases for gene integration and nucleases for marker gene deletion. *BMC Biotechnology*. **15**: 93.
- Nguyen L.V., Seok H.-Y., Woo D.-H., Lee S.-Y. & Moon Y.-H. 2018. Overexpression of the DEAD-Box RNA Helicase Gene AtRH17 Confers Tolerance to Salt Stress in Arabidopsis. *Int J Mol Sci*. **19**: .
- Nishizawa-Yokoi A., Cermak T., Hoshino T., Sugimoto K., Saika H., Mori A., Osakabe K., Hamada M., Katayose Y., Starker C., Voytas D.F. & Toki S. 2016. A Defect in DNA Ligase4 Enhances the Frequency of TALEN-Mediated Targeted Mutagenesis in Rice. *Plant Physiology*. **170**: 653–666.
- O'Malley R.C., Barragan C.C. & Ecker J.R. 2015. A User's Guide to the Arabidopsis T-DNA Insertional Mutant Collections. *Methods Mol Biol*. **1284**: 323–342.
- Parinov S., Sevugan M., Ye D., Yang W.-C., Kumaran M. & Sundaresan V. 1999. Analysis of Flanking Sequences from Dissociation Insertion Lines: A Database for Reverse Genetics in Arabidopsis. *The Plant Cell*. **11**: 2263–2270.
- Pauwels L., De Clercq R., Goossens J., Iñigo S., Williams C., Ron M., Britt A. & Goossens A. 2018. A Dual sgRNA Approach for Functional Genomics in Arabidopsis thaliana. *G3 Genes|Genomes|Genetics*. **8**: 2603–2615.

- Peterson P.A. 1970. The En mutable system in maize. *Theoret. Appl. Genetics*. **40**: 367–377.
- Pino M.-T., Skinner J.S., Park E.-J., Jeknić Z., Hayes P.M., Thomashow M.F. & Chen T.H.H. 2007. Use of a stress inducible promoter to drive ectopic AtCBF expression improves potato freezing tolerance while minimizing negative effects on tuber yield. *Plant Biotechnology Journal*. **5**: 591–604.
- Puchta H. 2005. The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J Exp Bot*. **56**: 1–14.
- Puchta H., Dujon B. & Hohn B. 1993. Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Research*. **21**: 5034–5040.
- Qi L.S., Larson M.H., Gilbert L.A., Doudna J.A., Weissman J.S., Arkin A.P. & Lim W.A. 2013. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*. **152**: 1173–1183.
- Qin J.Y., Zhang L., Clift K.L., Huler I., Xiang A.P., Ren B.-Z. & Lahn B.T. 2010. Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. *PLoS One*. **5**: .
- Qu S., Desai A., Wing R. & Sundaresan V. 2008. A Versatile Transposon-Based Activation Tag Vector System for Functional Genomics in Cereals and Other Monocot Plants. *Plant Physiology*. **146**: 189–199.
- Sallaud C., Gay C., Larmande P., Bès M., Piffanelli P., Piégu B., Droc G., Regad F., Bourgeois E., Meynard D., Périn C., Sabau X., Ghesquière A., Glaszmann J.C., Delseny M. & Guiderdoni E. 2004. High throughput T-DNA insertion mutagenesis in rice: a first step towards in silico reverse genetics. *The Plant Journal*. **39**: 450–464.
- Sanger F. & Coulson A.R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol*. **94**: 441–448.
- Sanger F., Nicklen S. & Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. **74**: 5463–5467.
- Schachtsiek J., Hussain T., Azzouhri K., Kayser O. & Stehle F. 2019. Virus-induced gene silencing (VIGS) in *Cannabis sativa* L. *Plant Methods*. **15**: 157.
- Sentenac H., Bonneaud N., Minet M., Lacroute F., Salmon J.M., Gaymard F. & Grignon C. 1992. Cloning and expression in yeast of a plant potassium ion transport system. *Science*. **256**: 663–665.
- Sessions A., Burke E., Presting G., Aux G., McElver J., Patton D., Dietrich B., Ho P., Bacwaden J., Ko C., Clarke J.D., Cotton D., Bullis D., Snell J., Miguel T., Hutchison D., Kimmerly B., Mitzel T., Katagiri F., Glazebrook J., Law M. & Goff S.A. 2002. A High-Throughput Arabidopsis Reverse Genetics System. *The Plant Cell*. **14**: 2985–2994.
- Shimizu Y., Bhakta M.S. & Segal D.J. 2009. Restricted spacer tolerance of a zinc finger nuclease with a six amino acid linker. *Bioorg Med Chem Lett*. **19**: 3970–3972.
- Şimşek Ö. & Aka Kacar Y. 2011. Discovery of mutations with TILLING and ECOTILLING in plant genomes. *Scientific Research and Essays*. **5**: 3799–3802.



- Smith J., Bibikova M., Whitby F.G., Reddy A.R., Chandrasegaran S. & Carroll D. 2000. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* **28**: 3361–3369.
- Speulman E., van Asperen R., van der Laak J., Stiekema W.J. & Pereira A. 2000. Target selected insertional mutagenesis on chromosome IV of Arabidopsis using the En-I transposon system. *Journal of Biotechnology.* **78**: 301–312.
- Svitashev S., Young J.K., Schwartz C., Gao H., Falco S.C. & Cigan A.M. 2015. Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol.* **169**: 931–945.
- Tang L., Mao B., Li Y., Lv Q., Zhang L., Chen C., He H., Wang W., Zeng X., Shao Y., Pan Y., Hu Y., Peng Y., Fu X., Li H., Xia S. & Zhao B. 2017. Knockout of OsNramp5 using the CRISPR/Cas9 system produces low Cd-accumulating indica rice without compromising yield. *Sci Rep.* **7**: 14438.
- Tang X., Liu G., Zhou J., Ren Q., You Q., Tian L., Xin X., Zhong Z., Liu B., Zheng X., Zhang D., Malzahn A., Gong Z., Qi Y., Zhang T. & Zhang Y. 2018. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biology.* **19**: 84.
- Tenea G.N. & Burlibasa L. 2012. RNAi Towards Functional Genomics Studies. *Functional Genomics.* .
- Till B.J. 2004. Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Research.* **32**: 2632–2641.
- Till B.J., Reynolds S.H., Weil C., Springer N., Burtner C., Young K., Bowers E., Codomo C.A., Enns L.C., Odden A.R., Greene E.A., Comai L. & Henikoff S. 2004. Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.* **4**: 12.
- Tovar-Aguilar A., Sánchez-Elizondo K., Rodríguez-Rodríguez A., González-Jaime M., Patino-Lopez G., Pérez V., Badillo-Corona J. & Durán-Figueroa N. 2019. Expression Pattern of Plant miRNAs by Classical Transcriptional Fusion Constructs: Methods and Protocols. pp. 175–185. *Methods in molecular biology* (Clifton, N.J.),
- Urnov F.D., Rebar E.J., Holmes M.C., Zhang H.S. & Gregory P.D. 2010. Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics.* **11**: 636–646.
- Warthmann N., Chen H., Ossowski S., Weigel D. & Hervé P. 2008. Highly Specific Gene Silencing by Artificial miRNAs in Rice. *PLOS ONE.* **3**: e1829.
- Wen Z., Yao L., Wan R., Li Z., Liu C. & Wang X. 2015. Ectopic Expression in Arabidopsis thaliana of an NB-ARC Encoding Putative Disease Resistance Gene from Wild Chinese Vitis pseudoreticulata Enhances Resistance to Phytopathogenic Fungi and Bacteria. *Front. Plant Sci.* **6**: .
- Wesley S.V., Helliwell C.A., Smith N.A., Wang M., Rouse D.T., Liu Q., Gooding P.S., Singh S.P., Abbott D., Stoutjesdijk P.A., Robinson S.P., Gleave A.P., Green A.G. & Waterhouse P.M. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant Journal.* **27**: 581–590.
- Wojcik A.M. 2020. Research Tools for the Functional Genomics of Plant miRNAs During Zygotic and Somatic Embryogenesis. *Int. J. Mol. Sci.* **21**: 4969.

- Woody S.T., Austin-Phillips S., Amasino R.M. & Krysan P.J. 2007. The WiscDsLox T-DNA collection: an arabidopsis community resource generated by using an improved high-throughput T-DNA sequencing pipeline. *J Plant Res.* **120**: 157–165.
- Xiao Y.-L., Redman J.C., Monaghan E.L., Zhuang J., Underwood B.A., Moskal W.A., Wang W., Wu H.C. & Town C.D. 2010. High throughput generation of promoter reporter (GFP) transgenic lines of low expressing genes in Arabidopsis and analysis of their expression patterns. *Plant Methods.* **6**: 18.
- Yao Y., Ni Z., Du J., Han Z., Chen Y., Zhang Q. & Sun Q. 2009. Ectopic Overexpression of Wheat Adenosine Diphosphate-ribosylation Factor, TaARF, Increases Growth Rate in Arabidopsis. *Journal of Integrative Plant Biology.* **51**: 35–44.
- Zhang N., Zhang D., Chen S.L., Gong B.-Q., Guo Y., Xu L., Zhang X.-N. & Li J.-F. 2018. Engineering Artificial MicroRNAs for Multiplex Gene Silencing and Simplified Transgenic Screen. *Plant Physiology.* **178**: 989–1001.
- Zhao M., Morohashi K., Hatlestad G., Grotewold E. & Lloyd A. 2008. The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. *Development.* **135**: 1991–1999.